

Towards the Prediction of Mortality in Intensive Care Units Patients: A Simple Correspondence Analysis Approach

Erika Severeyn¹, Miguel Altuve¹, Francisco Ng², Carlos Lollett¹, Sara Wong¹

¹Grupo de Bioingeniería y Biofísica Aplicada, Universidad Simón Bolívar, Caracas, Venezuela
²St Paul's Hospital, Vancouver BC, Canada

Abstract

In the setting of the PhysioNet/CinC Challenge 2012 Event 1, a new method to predict in hospital mortality in the Intensive Care Units (ICU) is proposed. The predictor, retrieved by Simple Correspondence Analysis (SCA), is based on a combination of clinical and laboratory data with more traditional score systems such as APACHE-II and SAPS-II. Information from records out of 12000 ICU patients was equally divided in three sets: A, B and C. Up to 37 variables were recorded during the first 48h after admission to the ICU. Using Set A, SCA was applied to select the variables most related to patients mortality from their hospitalizations. The proposed predictor combines these variables using the traditional APACHE II and SAPS II scores. SCA results show that variables such as creatinine, urine output, bilirubin and mechanical ventilation support were capable to discriminate between patients who survive or do not survive their ICU stays. Using these variables, the prediction method provides a SCORE1=43.50% using set A, SCORE1=42.25% using set B and SCORE1=42.73% using set C, where SCORE1 is defined as $\min(\text{sensitivity}, \text{positive predictivity})$. These results represent an improvement of 14% in SCORE1 when compared with traditional score SAPS-I (43.50% vs. 29.60%).

1. Introduction

The development of methods to estimate mortality in Intensive Care Unit (ICU) is not only motivated by the need to evaluate the quality of interventions provided in the critical care setting but also as an indicator of severity and outcome in acutely ill patients. Among the most popular mortality prediction models are the Acute Physiology and Chronic Health Evaluation II (APACHE II) proposed by Knaus et al. in 1985[1], the Simplified Acute Physiology Score II (SAPS II) proposed by Le Gall et al. in 1993[2] and the Sequential Organ Failure Assessment score (SOFA), proposed by Vincent et al. in 1996 [3]. These models are commonly used in ICU to

predict mortality, however, their accuracy depend on the pathological disease. For example, SAPS II and APACHE II show a poor mortality prediction in head injury patients [4].

Glasgow Coma Scale (GCS) is used to assess patient consciousness and many mortality prediction models, such as SAPS II, SOFA and APACHE II, include the component GCS. Nevertheless, the GCS by itself cannot predict the probability of mortality [4].

In this work, a new method to predict in hospital mortality in the Intensive Care Units (ICU) is proposed. The method is based on a combination of clinical and laboratory data with combined scores from APACHE II and SAPS II. Then, using Simple Correspondence Analysis (SCA), variables most related to in hospital death are exploited in the proposed predictor.

The rest of this work is organized as follows. In the next section, the database and the method to predict mortality in ICU, are presented. In section 3, SCA results as well as in hospital death predictions are exposed. Finally, conclusions and future works are outlined in section 4.

2. Method

2.1. Database

The information consists of records out of 12000 ICU patients, lasted at least 48 hours in the ICU. Records were divided in three sets: A, B and C, each one consisting of 4000 records. Set A was used to develop the predictor whereas sets B and C were used for validation purpose. Up to 41 variables were recorded once, more than once or not at all, during the first 48 hours after admission to the ICU. These variables were divided into three groups: general descriptors, outcomes related descriptors and time series [5].

General descriptors mainly were defined as age (AGE), gender (GEN), height (HEI), ICU type (ICU) and weight (WEI). These descriptors were collected when the patient was admitted into the ICU and they appear at the beginning of each record.

Outcome related descriptors were defined as SAPS score, SOFA score, length of stay in hospital (LOS), number of days between admission and death (SUR) and in hospital death (IHD-Survive, Dead). These descriptors were available only for training set A.

Times series variables are shown in Table 1. These variables were recorded in chronological order within each record.

Table 1. Time series variables recorded in the ICU

| Variables | Description |
|-----------|---|
| ALB | Albumin (g/dL) |
| ALP | Alkaline phosphatase (IU/L) |
| ALT | Alanine transaminase (IU/L) |
| AST | Aspartate transaminase (IU/L) |
| BIL | Bilirubin (mg/dL) |
| BUN | Blood urea nitrogen (mg/dL) |
| CHO | Cholesterol (mg/dL) |
| CREA | Serum creatinine (mg/dL) |
| DBP | Invasive diastolic arterial blood pressure (mmHg) |
| FIO | Fractional inspired O2 (0-1) |
| GCS | Glasgow Coma Score (3-15) |
| GLU | Serum glucose (mg/dL) |
| HCO | Serum bicarbonate (mmol/L) |
| HCT | Hematocrit (%) |
| HR | Heart rate (bpm) |
| K | Serum potassium (mEq/L) |
| LAC | Lactate (mmol/L) |
| MG | Serum magnesium (mmol/L) |
| MAP | Invasive mean arterial blood pressure (mmHg) |
| MEVE | Mechanical ventilation respiration |
| NA | Serum sodium (mEq/L) |
| NBP | Non-invasive diastolic arterial blood pressure (mmHg) |
| NAP | Non-invasive mean arterial blood pressure (mmHg) |
| NSP | Non-invasive systolic arterial blood pressure (mmHg) |
| PCO | partial pressure of arterial CO2 (mmHg) |
| PO2 | Partial pressure of arterial O2 (mmHg) |
| PH | Arterial pH (0-14) |
| PLA | cells/nL |
| RRA | Respiration rate (bpm) |
| SO2 | O2 saturation in hemoglobin (%) |
| SBP | Invasive systolic arterial blood pressure (mmHg) |
| TEM | Temperature (°C) |
| TRI | Troponin-I (µg/L) |
| TRT | Troponin-T (µg/L) |
| URI | Urine output (mL) |
| WBC | White blood cell count (cells/nL) |
| WEI | kg |

2.2. ICU mortality predictor

SCA is commonly applied to datasets in which a group of individuals is described by groups of variables, in order to find common structures existing in these groups of individual [6].

In this work, a SCA was used to identify relationships between each variable and the outcome related descriptor: ‘in-hospital death’. Level of association among each patient including general descriptors and clinical data with time series within their 48 hours window while admitted to the ICU were analyzed with the outcome related descriptor.

As can be seen in Figure 1, the proposed predictor combines the variables more related to ‘in hospital death’ and ‘survive’, using the scores from APACHE II and SAPS II methods. The output of the predictor is 0 if the patient survives and 1 if he does not survive. The predictor performance was evaluated using sensitivity (Se), positive predictivity (+P) and SCORE1= $\min(\text{Se}, +P)$.

In the training phase, a Cutoff Point (CP) associated with the detector score mortality event, was titrated in order to obtain the best SCORE1 (figure 1). The CP is the decision point in the score which will be determined if the person will survive; in this case, subjects with a score higher than CP will not likely survive.

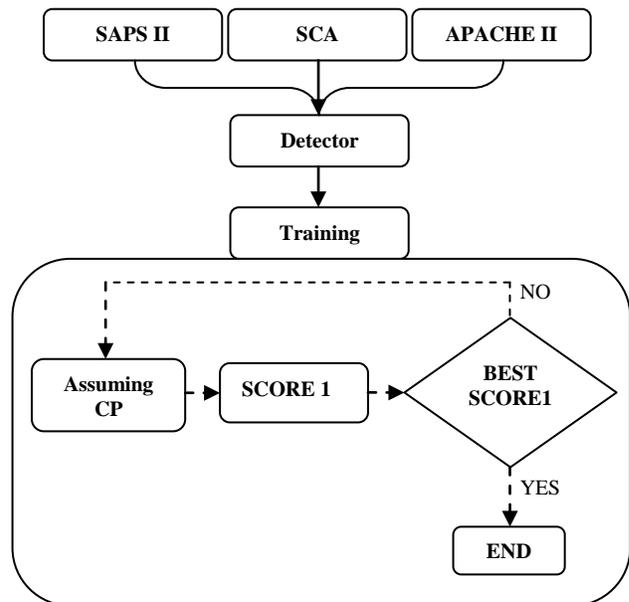


Figure 1. Methodology used to develop the predictor.

3. Results

3.1. Simple correspondence analysis

Figure 2 shows the SCA for set A and Table 2 shows the contributions of the methods in two axis of the SCA, for set A. Variables with major contributions in the first axis are: BUN (25.05%), CREA (30.37%), URI (9.25%)

and SOFA (9.14%). In the second axis, the major contributions came from: LOS (23.9%), SOFA (10.97%), CREA (22.58%) and BUN (12.04%). Axis 2 is divided in two regions: subjects who did not use mechanical ventilation support (upper quadrants) and subjects who used mechanical ventilation support (lower quadrants). Subjects who required mechanical ventilation support (MECHVENT) are more related to variable DEAD and subjects who did not use mechanical ventilation (NOT

MECHVENT) are more related to variable SURVIVE.

SCA results show that variables such as CREA, BUN, BIL, SOFA, MECHVENT and SAPS are more related to variable DEAD. Variables URI, GCS, PLA, NOT MECHVENT and HCO are more related to variable SURVIVE, indicating that those variables are able to discriminate between patients who survive or do not survive their ICU stays.

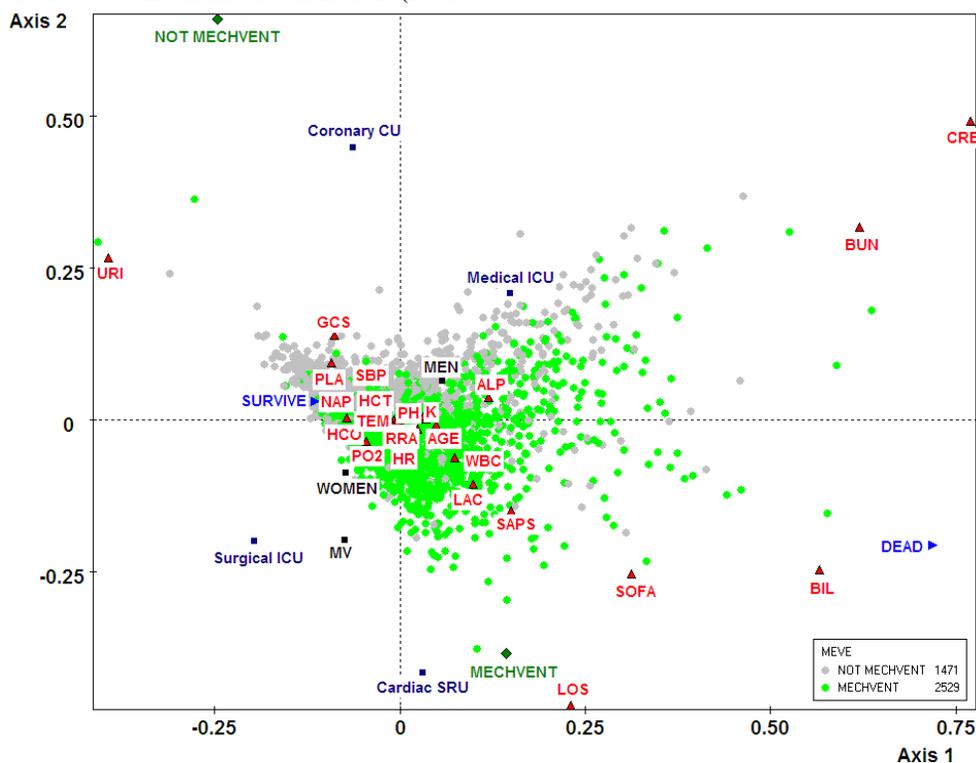


Figure 2. SCA for set A

Table 2. Contributions of the variables in the SCA.

| Variable | Axis 1 | Axis 2 |
|----------|--------|--------|
| ALP | 0,74 | 0,13 |
| BIL | 7,29 | 2,50 |
| BUN | 25,05 | 12,04 |
| CRE | 30,37 | 22,58 |
| GCS | 1,46 | 6,74 |
| URI | 9,25 | 7,74 |
| SAPS | 3,48 | 6,13 |
| SOFA | 9,14 | 10,97 |
| LOS | 3,16 | 23,90 |

3.2. Predictor evaluation

The proposed predictor was designed based on information from the SCA and combines the scores of those variables from APACHE II and SAPS II.

Variables used and scores assigned to each one, for the total score of the predictor proposed, are shown in Table 3. Our proposed predictor is a scoring system based on

values obtained in Table 3 and an optimal cut-off point as a decision-making parameter.

Using these variables and with an optimal CP=24.25, the predictor provides: Se=41.15%, +P=41.01% using set A, and SCORE1=41.02% using set B. Moreover, with a CP=24.15, the predictor shows: Se=41.87%, +P=40.84% using set A, and SCORE1=41.90% using set B. In the phase 2 with an optimal CP=24.5 provides: Se=45.31%, +P=43.50% using Set A, a SCORE1=42.25% using Set B and a SCORE1=42.73 using set C.

4. Conclusions and future works

In this work, a new method to predict mortality in patients admitted to intensive care units is proposed. Firstly, a SCA was used to find relationships among clinical values and the variable 'in hospital death'. This technique show that variables such as creatinine, urine output, BUN, bilirubin, SOFA score, SAPS score, GCS and mechanical ventilation support, are able to discriminate between patients who survive or do not survive their ICU stays. Afterwards, a combination of

scores from SAPS II and APACHE II are used to develop the predictor. The proposed method improves SCORE1 in 14% when compared with traditional score such as SAPS-I (43.5% vs. 29.60%). Additionally, this predictor is highly reproducible as SCORE1 was comparable in all

three sets.

Future work is related to the optimization of the punctuation in each variable through an evolutionary algorithm, in order to improve the predictor performance.

Table 3. Variables punctuation for the proposed detector.

| Variable | Punctuation | | | | |
|----------------|---------------------|----------------------------|--------------------------|--------------------------|-----------------------|
| | 0 | 1 | 2 | 3 | 4 |
| TEMP | 38.5>T≥36 | 36>T≥34 | 34>T≥32 | 41>T≥39 32>T≥30 | T≥41 30>T>0 |
| NAP | 110>NAP≥70 | | 130>NAP≥110 70>NAP≥50 | 160>NAP≥130 | NAP≥160 0<NAP<50 |
| HR | 110>HR≥70 | | 140>HR≥110 70>HR≥55 | 180>HR≥140 55>HR≥40 | HR≥180 40>HR>0 |
| MEVE | Do not use | | | | Use |
| FIO ≥0.5 (0-1) | PO ₂ ≥70 | 70>PO ₂ ≥61 | | 61>PO ₂ ≥55 | 55>PO ₂ >0 |
| PH | 7.5>pH≥7.33 | 7.6>pH≥7.5 7.33>pH≥7.25 | 7.25>pH≥7.15 | 7.7>pH≥7.6 | pH≥7.7 0<pH<7.15 |
| NA | 150>Na≥130 | 155>Na≥150 | 160>Na≥155 130>Na≥120 | 180>Na≥160 120>Na≥110 | Na≥180 0<Na<110 |
| K | 5.5>K≥3.5 | 6>K≥5.5 <6 3.5>K≥3 | 3>K≥2.5 | 7>K≥6 | K≥7 0<K<2.5 |
| CRE | 1.5>CREA≥0.6 | 2>CREA≥1.5 | 0<CREA <0.6 | 3.5>CREA≥2 | CREA≥3.5 |
| HCT | 46>HCT≥30 | 50>HCT≥46 | 60>HCT≥50 30>HCT>20 | | HCT≥60 0<HCT<20 |
| WBC | 15>WBC≥3 | 20>WBC≥15 | 40>WBC≥20 3>WBC≥1 | | WBC≥40 0<WBC<1 |
| AGE | 0 <AGE<40 | 60>AGE≥40 | 70>AGE≥60 | 74>AGE≥70 | AGE≥74 |
| PLA | PLA≥120 | 120>PLA≥81 | 81>PLA≥51 | 51>PLA ≥21 | 21≥PLA>0 |
| BUN | 0<BUN<10 | | 30>BUN≥10 | | BUN≥30 |
| RRA | 25>RR≥12 | 35>RR≥25 12>RR≥10 | 10>RR≥6 | 50>RR≥35 | RR≥50 0<RR<6 |
| SBP | 200>SBP≥100 | SBP≥200 | | 100>SBP≥70 | 0<SBP<70 |
| BIL | BIL≤21 | 61≥BIL>21 | 121≥BIL>61 | 240≥BIL>121 | BIL>240 |
| HCO | HCO≥19 | | 19>HCO≥15 | | 0<HCO<15 |
| URI | URI≥1 | | 1>URI≥0.5 | | 0<URI<0.5 |
| GCS | | | 15-GCS | | |

References

- [1] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHEII: a severity of disease classification system. Crit Care Med 1985;13:818–29..
- [2] Le Gall JR, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. JAMA 1993;270:2957–63.
- [3] Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, Reinhart CK, Suter PM, Thijs LG. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. Intensive Care Med 1996; 22(7):707-10.
- [4] Jennet B, Teasdale G, Braakman R, Minderhoud J, Heiden J, Kurze T. Prognosis of patients with severe head injury. Neurosurgery 1979;4:283–9.
- [5] Predicting Mortality of ICU patients. Available in <http://physionet.org/challenge/2012/>. 2012
- [6] Lebert L., Morineau A, Piron M. Statistique Exploratoire Multidimensionnelle. Paris. Dunod. 3rd edition. 2000: 344-346.

Address for correspondence.

Erika Severeyn

Universidad Simón Bolívar, Caracas, Venezuela, 89000.

severeynrika@usb.ve